

CV703: Week 5 Lab

Object Recognition using Vision Transformer

The following tasks will be performed in the lab today. Additional details, related to this lab, are provided in the Jupyter notebook. The objective here is to complete the tasks below. Further, the aim is to understand the entire notebook code after completing the tasks.

Task- Finetune Vision transformer on Stanford Dogs dataset.

1. Load Stanford Dogs dataset.
2. Load ViT.
3. Perform ViT fine-tuning on training set of Dogs dataset.
4. Test ViT performance on Dogs test set.
5. Check network performance by increasing use_top_n_heads.
6. Check network performance by increasing number of epochs.